Inspection of mechanical assemblies based on 3D Deep Learning segmentation

Assya Boughrara, Igor Jovančević, Jean-José Orteu, Mathieu Belloc

Inspection of mechanical assemblies based on 3D Deep Learning segmentation

Assya Boughrara[a], Igor Jovančević[b], Jean-José Orteu[a], and Mathieu Belloc[c]

[a]Institut Clément Ader (ICA) ; Université de Toulouse ; CNRS, IMT Mines Albi, INSA, UPS, ISAE ; Campus Jarlard, 81013 Albi, France
[b]Faculty of Natural Sciences and Mathematics, University of Montenegro, Cetinjska 2, 81000 Podgorica, Montenegro
[c]Diota, 201 Pierre and Marie Curie Street, 31670 Labège, France

ABSTRACT

Our research work is being carried out within the framework of the joint research laboratory "Inspection 4.0" between IMT Mines Albi/ICA and the company *Diota* specialized in the development of numerical tools for Industry 4.0. In this work, we are focused on conformity control of complex aeronautical mechanical assemblies, typically an aircraft engine at the end or in the middle of the assembly process. A 3D scanner carried by a robot arm provides acquisitions of 3D point clouds which are further processed by deep segmentation networks. Computer-Aided Design (CAD) model of the mechanical assembly to be inspected is available, which is an important asset of our approach. Our deep learning models are trained on synthetic and simulated data, generated from the CAD models. This research is a continuation of the work presented at the QCAV'2021 conference [1].

Keywords: aeronautics, robotized inspection, quality control, CAD model, 3D deep learning, 3D point cloud, segmentation.

1. INTRODUCTION

Automatic control by artificial vision is a way to relieve operators of repetitive visual inspection tasks and to ensure a faster and better quality control, by also providing traceability of control operations. Providing such control tools is one of the objectives of the company *Diota*, which is an industrial partner in this work. Specialized in the development of digital tools for Industry 4.0, *Diota* offers innovative solutions that exploit PLM (product lifecycle management) data, augmented reality technology and non-contact inspection technics.

This work is part of the development of "Digital-based Robotics" tools. The Computer-Aided Design (CAD) model is used as a reference for this inspection process. It is also used as a guide for the inspection and it allows us to know a priori which mechanical part is being inspected.

The task is to establish the conformity of the inspected part based on the CAD. A conformity is defined by the actual presence of the expected mechanical part within the assembly. A proposed inspection module must be able to differentiate a compliant case from any other possible situation that may arise : (1) The expected part is absent; (2) Another part than the expected one is mounted (mounting error). Examples of three possible scenarios for the parts of type "support", are shown in figure 1.

The control end effector offers great flexibility and inspection capability. It is composed of several sensors offering different fields of view, which allows several functionalities (see figure 2).

In the work presented at QCAV'2021 [1], we addressed the problem with a pure classification approach. However, this approach showed weaknesses for missing support detection. Therefore, we have explored the semantic segmentation approach, more precisely the part segmentation approach, which will be presented in this paper.

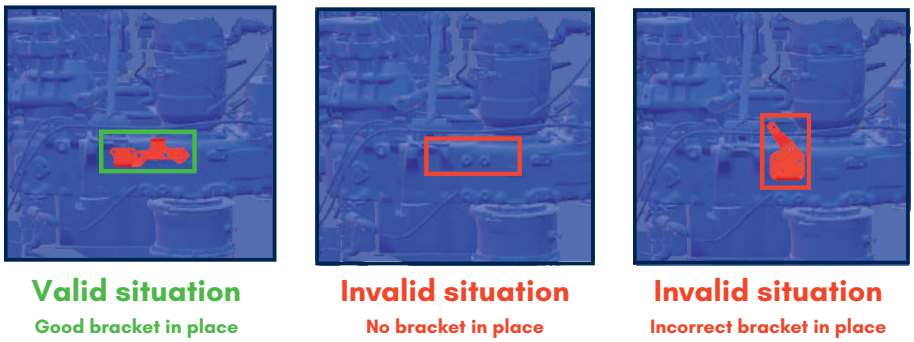


Figure 1: Examples of the different possible scenarios

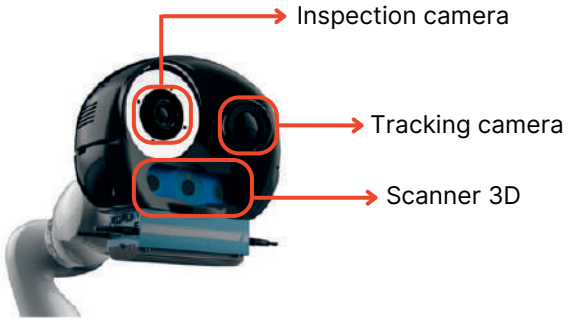


Figure 2: Robotic end effector developed by *Diota*, with 3 mounted sensors : an inspection camera (small field of view), a tracking camera (wide field of view) and a 3D acquisition sensor Ensenso N35.

2. PREVIOUS WORKS

The rise of artificial intelligence (AI) has benefited many fields. This is the case, for example, for autonomous cars, but more recently for the aerospace sector as well. Among major tasks which could involve AI, Airbus, and more generally the various aerospace manufacturing players, focus on anomaly detection and decision making.

In production or aeronautical maintenance, the vast majority of visual inspection tasks are aimed at finding defects or anomalies. These are typically long tasks requiring a great concentration of the operator. Furthermore, these detections are subject to human errors. With ever-increasing air traffic and high demands on maintenance personnel to meet commercial schedules, the pressure on inspection operations is becoming higher, which increases the risk of errors [2, 3].

In aeronautics, the first robotic solutions focused on fuselage inspection, the external surface of the aircraft, with a robot crawling on the aircraft. For example [4] is based on ANDI (Automated NonDestructive Inspector) and CIMP (Crown Inspection Mobile Platform). The goal was to detect cracks and corrosion on the aircraft fuselage. These works were the pioneers of remote 3D stereoscopic visual inspection. NASA's work [5], with the MACS robot, also belongs to the family of crawling robots. The 2010s were marked by the launch of a collaborative mobile robot with wheels named Air-Cobot, implemented for aeronautical inspection. Air-Cobot is capable of moving safely around an aircraft in an environment that contains obstacles to avoid. This application has led to many works [6, 7, 8]. This robot is equipped with a Pan-Tilt-Zoom (PTZ) camera and a 3D scanner, both used for inspection. There are several works on inspection based on the analysis of 2D images from PTZ camera acquisitions [9,10, 11]. Other works have focused on the exploitation of 3D data, also for visual inspection [12, 13].

In many industries, inspection of complex products with variable configurations by conventional manual methods can be cumbersome and limited. Time consuming and costly, these controls can be incomplete, prone to errors and often lack traceability. In this context, our partner *Diota*, has launched a project on the inspection of complex aeronautical assemblies in the production phase, based on exploiting 2D images or 3D point clouds [14,

15, 16]. This work is a continuation of these endeavours, focusing on the exploitation of deep learning techniques to analyze 3D point clouds. In preliminary experiment [1], we have chosen PointCNN [17], the network which has proven to be the best in our classification tasks. We decide to use this network for a part-based segmentation approach.

3. DATASETS PROCESSING

For our inspection problem, we strive to obtain 3D deep learning models able to infer on real data without having seen any during the training. For this, we choose to perform our training only on synthetic data. Here our synthetic data are synthetic 3D point clouds generated from a CAD model.

3.1 Available data

CAD provides different levels of information: (1) support-fusion (cf. figure 3a): support (in red), with the elements of its close context like screws or clamps (in green); (2) support-fusion with its large context (in blue): corresponds to the wider context in which the support-fusion is integrated, for example the pipes and cables surrounding the support-fusion (cf. figure 3b).

(a) (b)

Figure 3: Visualization of the synthetic scene from the same point of view. (a) support-fusion only, (b) support-fusion and its large context (in blue).

Further, figure 4 shows synthetic rendering for each real acquisition.

Figure 4: The first row represents the 2D images taken by the camera with a wide field of view. The second row represents the synthetic renderings corresponding respectively to the image acquisitions from the first row.

3.2 Datasets generation

This process is essentially based on the combination of a Fibonacci sphere centered in the CAD of the inspected part (typically support) and a Z-buffer method for surface rendering. This approach takes into account occlusion and self-occlusion, while ensuring a realistic rendering very similar to the one obtained with the scanner. In order to simulate real data acquisition conditions, the virtual scanner is positionned at 60 cm from the centroid of the support. The number of viewpoints around the support is a parameter named N. N determines the discretization of the Fibonacci sphere, and it was defined by preliminary experiments. Different forms of 3D

data-augmentation can be integrated to the generation of synthetic data. We apply the *Farthest Point Sampling* (FPS) [18], a dowsampling method, in order to keep a fixed number of uniformly distributed points per synthetic acquisition. A summary diagram of the synthetic data generation pipeline is presented in figure 5.

Data augmentation should allow the *deep learning* model to become invariant to the variabilities to which it may be exposed in the inference phase. Indeed, industrial CAD data can be qualified as "perfect" data, without any noise or occultation by some external *artefacts* such as cables or protective plastic not present in CAD. Therefore, we implemented data augmentation methods specific to our problem, adding, during the generation of synthetic clouds:

- **Gaussian noise to the point positions**: simulating the acquisition noise of the scanner
- **Artefacts**: simulating a partial occultation, like the one caused by a cable.

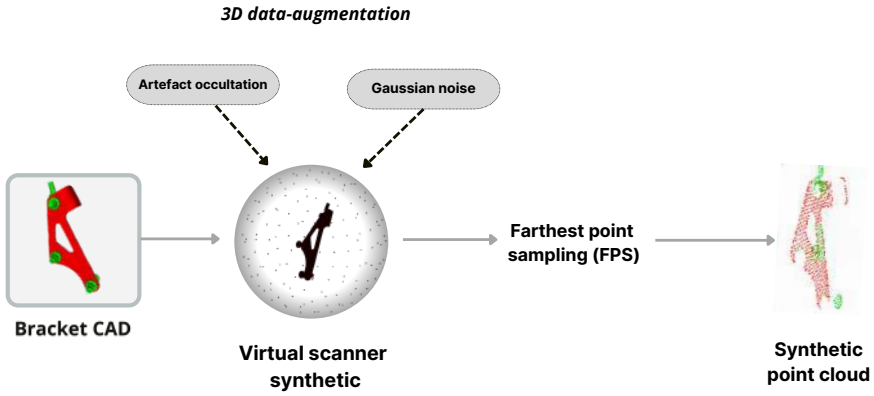
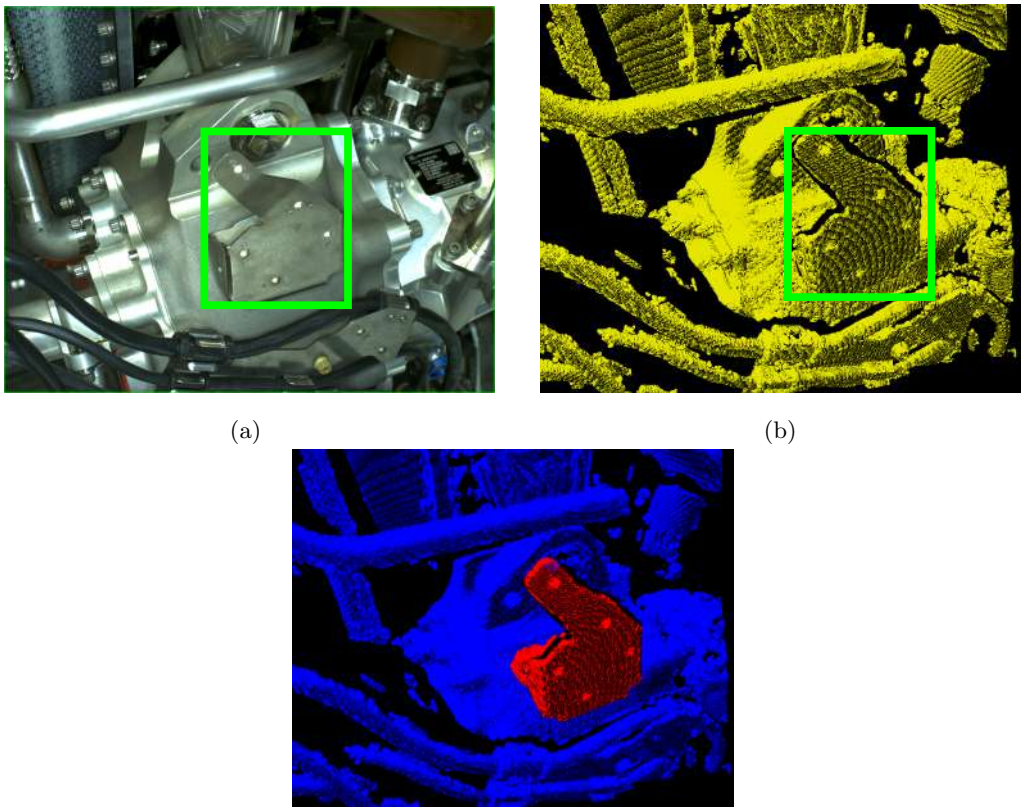


Figure 5: Complete description of our synthetic data generation process

During the synthetic data generation we add the large context. After the generation of our synthetic cloud, we will provide two labels (two segmentation classes) to segmentation models: (1) support-fusion (cf. figure 3a), 2) large context (cf. figure 3b).

3.3 Automatic annotation

For the evaluation of our segmentation models we need the point-level labels of our real point cloud acquisitions. So we exploit the digital model to segment the support-fusion in the scanner point cloud acquisition. For this purpose, we developed a module that aligns the digital CAD model to the 3D point cloud acquisition. Thus allowing to separate the support-fusion from the background (context) in the point cloud. An example shows the result of automatic annotation on a real cloud in figure 6. It is important to note that we used the alignment module only to generate labels on already almost alligned examples. It should be stressed that this method, which is very sensitive to pose shifts and variations, could not replace our segmentation module for the detection of missing supports.



(a) (b)

(c)

Figure 6: **Automatic annotation :** (a) 2D image corresponding to the acquisition, with the bounding box around the support 21 in green; (b) the 3D point cloud provided by the scanner, with the bounding box around the support 21 in green; (c) completed labeling of the real acquisition by our automatic annotation module based on the synthetic model. In red the support and in blue the context (background).

4. EXPERIMENTS

Our previous experiments [1] allowed us to choose PointCNN as the best network on our data for a classification task. The knowledge extraction part of the network being common for classification and segmentation, we chose to use the same network for our part segmentation task.

4.1 Our model

For training, we generated 3D point clouds database using 3D CAD models from 61 different supports. To generate synthetic 3D point clouds from many different points of view, we have developed a z-buffer based method. Same as the scanner, this method takes into account self-occlusion. For each support with its surrounding (cf. figure 3b), we generate 40 different synthetic clouds and ignore the views where the background represents more than 80% of the points. We have two distinct labels: 0 for background and 1 for support. The distribution of the clouds is Train-set: 945; Validation-set: 117; Test-set: 150. After the training, PointCNN achieves 97.2% accuracy and 94% MIoU on the synthetic Testing-set. We have tested this model on 623 real acquisitions: 590 presence cases and 33 absence cases. The threshold (minimum) for missing case detection has been set to 80% of points segmented as background. Table 1 presents the results and the confusion matrix is given in table 2.

These results are very encouraging and confirm a good ability of a segmentation-by-parts approach, trained on synthetic data, to obtain good results on real data. However the false negative (FN) rate (absence recognized as a presence) greater than 0 is not acceptable.

Set	*Accuracy*	*Precision*	*Recall*	*F1-score*	Nb acq
Presence set	98.8%	100%	98.8%	99.4%	590
Absence set	96.9%	100%	96.9%	98.4%	33
Total	98.7%	98.8%	98.7%	98.7%	623

Table 1: Results on real point clouds with threshold 80%

label/pred	Missing	Presence	Nb acq.
Missing	32 (TP)	1 (FN)	33
Presence	7 (FP)	583 (TN)	590

Table 2: Confusion matrix after application of the 80% threshold for missing case detection.

After a close look into the acquisitions that caused errors, we could see that these acquisitions contain reflective areas that confused the scanner which resulted either in a lack of points, or in an extremely noisy acquisitions.

4.2 Evaluation

Several segmentation results on real acquisitions are presented in figure 7.

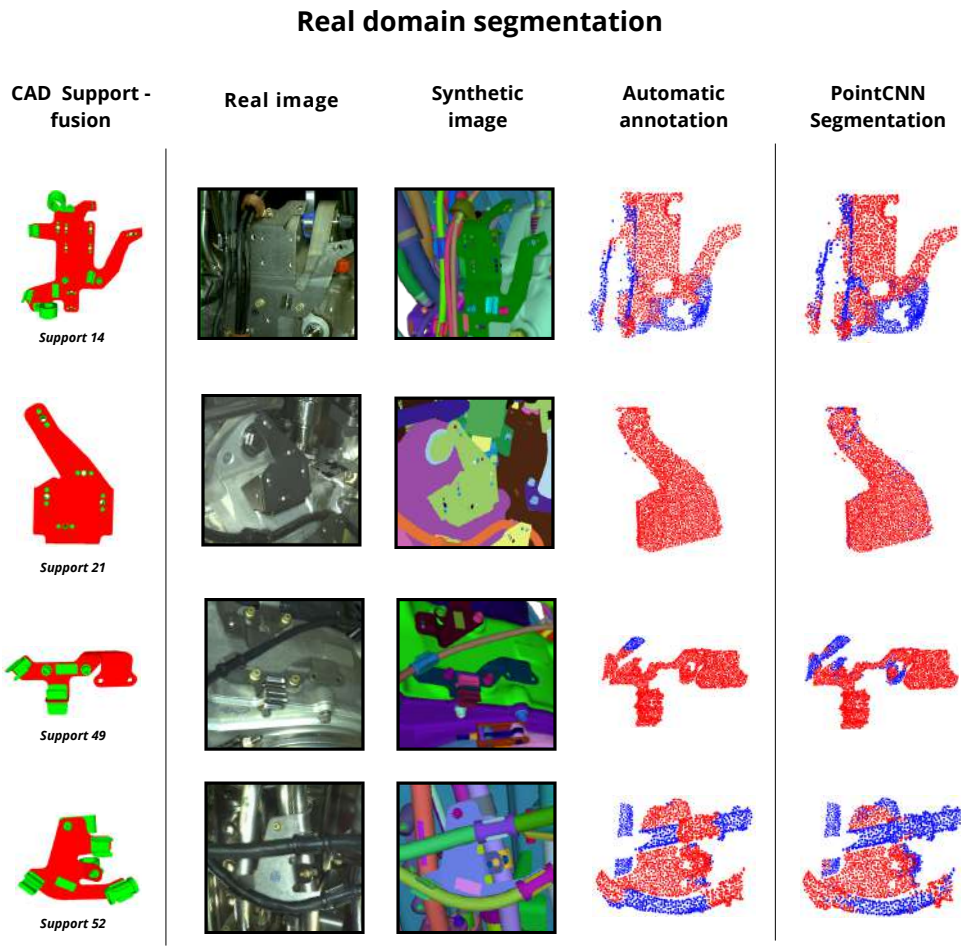


Figure 7: Results on real scanner acquisitions

On all of these visualizations, we see almost the same segmentation results provided by our PointCNN model, compared to those provided by our annotation module. As for support 21, despite an almost perfect segmentation, we find context points scattered in the middle of the support area. As for the support 52, we find a robustness to separate the support despite a dense context with many cables/pipes. This robustness to occlusion is interesting to note, despite some difficulties in discerning cable clamps in some places. We can confirm that synthetic learning allows, a relatively good robustness against cables and occultations on the real domain. For supports 49 and 14, the proposed segmentation is very close to that of the automatic annotation, with some regions wrongly segmented as context and vice versa.

Another problem on which we wanted to focus our attention is that of density variation. Indeed, real acquisitions are imperfect by definition, compared to purely synthetic acquisitions. It is therefore not uncommon to see density variations in the real acquisitions. It is important that, when faced with this type of example, our network is still able to identify the presence of the support. It is the case for the acquisition of the support 14, (cf. figure 8), where, in spite of great variations in density, the segmentation proposed by PointCNN manages to draw the contours of the support.

(a) *PointCNN* segmentation (b) Automatic annotation

Figure 8: **Density variation :** acquisition with large density variations for the support 14.

3D point cloud segmentation with 2 classes can be described as a point-wise classification problem. In our study, to obtain an evaluation by support, we calculate Mean Intersection over Union (MIoU) for each support. For each support, MIoU is expressed as shown in equation (1), with $type_i$ corresponding to the support i. IoU_j represents the ratio of the number of correctly classified points and the total number of classified points, for the j^{th} cloud of the support $i \in \{1...M\}$. n_i represents the number of clouds corresponding to support i.

$$MIoU(type_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} IoU_j \tag{1}$$

This metric allows to obtain an overall view by class. This overview considers all the present labels (two labels in our case) without giving importance to any label. Only the proportion of correctly classified points counts, regardless of their label. Here, $M = 61$ since there are 61 different supports that we consider in our assembly. For our evaluation we use as reference the automatic annotation presented in section 3.3. The *average MIoU* reaches 75% on the Presence set and 92% on the Absence set. These are good results. The less good performance on the Presence set is due to some small wrongly classified regions. An example is the support 21 in figure 7 where many context points are found in the middle of those associated with the support.

To improve these results we propose a post processing method based on radius neighbors to eliminate wrongly classified regions. This method can be described as weighted neighborhood average smoothing. This method was able to increase the *average MIoU* on the Presence set by 2%, reaching 77%. On the Absence set the performance did not change. It should be noted that even if smoothing presents better visual results, unfortunately it did not allow to decrease the FP or FN previously presented in table2.

5. CONCLUSION

In this article, we sought to define a protocol to automatically recognize the absence/presence of a mechanical assembly support. In order to detect the presence of support, we employed the known PointCNN segmentation network, which we trained with purely synthetic 3D point clouds. We trained the part-based segmentation network only on the clouds with present support. We trained the network to extract only the support from the cloud, i.e. to separate the support from its surrounding (context). From the segmentations provided by PointCNN, we further propose an inspection protocol to recognize the presence/absence of support. Thus, we set a threshold for absence detection based on the proportion of context points in the cloud. We chose the threshold of 80% which allows us to reach an Accuracy of 98.7% on our real data set. In terms of absolute values, we obtain 1 FN and 7 FP. Expectedly, we found that the network has proven to be less accurate with reflective areas. This is an objective of our further work. This article was the basis for our complete inspection module, which also contains support type recognition.

ACKNOWLEDGMENTS

This research work has been carried out within a PhD thesis co-funded by the French "Région Occitanie." We would like to thank the "Région Occitanie" for its financial support.

REFERENCES

[1] Boughrara, A., Jovančević, I., Ben Abdallah, H., Dolives, B., Belloc, M., and Orteu, J.-J., "Inspection of mechanical assemblies based on 3D Deep Learning approaches," in [*QCAV'2021 - 15th International Conference on Quality Control by Artificial Vision*], *Proceedings of SPIE* **11794**, 8 p. (May 2021).

[2] Max, D. and Graeber, R., "Human error in maintenance," *Aviation Psychology in Practice* **1**, 87–104 (1994).

[3] Drury, C., "Human reliability in civil aircraft inspection," tech. rep., STATE UNIV OF NEW YORK AT BUFFALO DEPT OF INDUSTRIAL ENGINEERING (2001).

[4] Davis, I. L. and Siegel, M., "Automated nondestructive inspector of aging aircraft," in [*Measurement Technology and Intelligent Instruments*], **2101**, 190–201, SPIE (1993).

[5] Bar-Cohen, Y., "Autonomous rapid inspection of aerospace structures," (Sep 1997).

[6] Futterlieb, M., Cadenat, V., and Sentenac, T., "A navigational framework combining visual servoing and spiral obstacle avoidance techniques," in [*2014 11th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*], **2**, 57–64, IEEE (2014).

[7] Frejaville, J., Larnier, S., and Vetault, S., "Localisation à partir de données laser d'un robot naviguant autour d'un avion," in [*Reconnaissance des Formes et l'Intelligence Artificielle*], (June 2016).

[8] Bauda, M.-A., Bazot, C., and Larnier, S., "Real-time ground marking analysis for safe trajectories of autonomous mobile robots," in [*2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM)*], 1–6 (2017).

[9] Jovančević, I., Larnier, S., Orteu, J.-J., and Sentenac, T., "Automated exterior inspection of an aircraft with a pan-tilt-zoom camera mounted on a mobile robot," *Journal of Electronic Imaging* **24**(6), 061110 (2015).

[10] Jovančević, I., Pham, H.-H., Orteu, J.-J., Gilblas, R., Harvent, J., Maurice, X., and Brèthes, L., "3d point cloud analysis for detection and characterization of defects on airplane exterior surface," *Journal of Nondestructive Evaluation* **36**, 74 (Oct 2017).

[11] Leiva, J. R., Villemot, T., Dangoumeau, G., Bauda, M.-A., and Larnier, S., "Automatic visual detection and verification of exterior aircraft elements," in [*2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM)*], 1–5, IEEE (2017).

[12] Jovančević, I., *Exterior inspection of an aircraft using a Pan-Tilt-Zoom camera and a 3D scanner moved by a mobile robot: 2D image processing and 3D point cloud analysis*, PhD thesis, Ecole des Mines d'Albi-Carmaux (2016).

[13] Bauda, M.-A., Grenwelge, A., and Larnier, S., "3d scanner positioning for aircraft surface inspection," in [*ERTS 2018*], (2018).

[14] Ben Abdallah, H., *Inspection d'assemblages aéronautiques par vision 2D/3D en exploitant la maquette numérique et la pose estimée en temps-réel*, PhD thesis, Université de Toulouse - IMT Mines Albi (Février 2020).

[15] Ben Abdallah, H., Orteu, J.-J., Dolives, B., and Jovančević, I., "3d point cloud analysis for automatic inspection of aeronautical mechanical assemblies," in [*Fourteenth International Conference on Quality Control by Artificial Vision*], **11172**, 209–217, SPIE (2019).

[16] Mikhailov, I., Jovancevic, I., Mokhtari, N. I., and Orteu, J.-J., "Classification using a 3D sensor in a structured industrial environment," *Journal of Electronic Imaging* **29**, 041008 (July 2020).

[17] Li, Y., Bu, R., Sun, M., and Chen, B., "Pointcnn," *CoRR* **abs/1801.07791** (2018).

[18] Kamousi, P., Lazard, S., Maheshwari, A., and Wuhrer, S., "Analysis of Farthest Point Sampling for Approximating Geodesics in a Graph," *CoRR* **abs/1311.4665** (2013).